Two types of inductive analogy by similarity

Kuipers, Theo A.F.

Published in:
Erkenntnis, vol. 21 (1), 63-87

THEO A. F. KUIPERS

TWO TYPES OF INDUCTIVE ANALOGY BY SIMILARITY

1. INTRODUCTION

In recent years some new work has appeared about the problem of integrating into systems of inductive probability analogy influences based on a distance function between the predicates, so-called analogy by similarity. Carnap's last (manuscript) attempts were published (Carnap (1980), Sections 16 and 17[1]), Niiniluoto (1981) formulated a weak and a strong 'measure'. Spohn (1981) commented on them and has introduced a new proposal.

None of these attempts seem to be very successful. Carnap restricts the discussion almost completely to 'sequences' of two predicates. As Spohn has pointed out, Niiniluoto's first measure violates positive instantial relevance and the second one violates Reichenbach's axiom. Both properties are certainly a prerequisite for inductive probability. In the second part of his paper, Spohn presented a system which has both properties and, in addition, it has regularity, initial symmetry between the predicates, some analogy influences and positive probabilities for contingent universal generalizations. But Spohn is well aware that his system is not symmetric with respect to (the order of) individuals and, more importantly, that it is very ad hoc, because it lacks a plausible way of introduction or interpretation.

There is also a paper by Costantini (1983) on the subject. We will comment on it in some detail in Section 3.

In this paper we will explore the idea, briefly described on earlier occasions (e.g. Kuipers (1981), Section 3), that there are two quite different ways in which analogy (by similarity) influences can be conceived. The one, called *instantial analogy*, is based on the idea that predicates can 'acquire analogy' on the basis of evidence. This intuition plays the dominant role in most publications. The other, called *existential analogy*, is based on the idea that the distance between the predicates can play a role in the prior distribution.

After a summary of the general background in Section 2 we will present

Erkenntnis **21** (1984) 63-87. 0165-0106/84/0211-0063 $02.50

in Section 3 a class of Carnap-like systems in which instantial analogy influences arise on the basis of 'virtual analogy instances' which are distributed on the basis of the available evidence. These systems have the properties (using our favourite terminology): initial (predicate) symmetry, instantial confirmation (positive instantial relevance), instantial convergence (Reichenbach's axiom). Moreover, they are closed, i.e. they lead to zero probability for contingent universal generalizations. Finally, they have the property called virtual analogy, which turns out to be a more acceptable alternative for a general principle of analogy than the one proposed by Carnap. From our set up it follows quite naturally that 'Carnap-analogy' is only guaranteed if the distance function is of a special kind.

But the systems developed do not have order indifference (exchangeability), although they do have some order convergence properties. This lack of order indifference will be defended in some detail.

In Section 4 we will introduce a class of open systems (systems with non-zero probabilities for universal generalizations) which have the properties of order indifference, instantial confirmation and convergence, and universal confirmation and convergence. Moreover, four conditions of increasing strength will be imposed upon the prior distribution on the basis of the distance function and all of them will lead to plausible properties for the posterior distribution. However, such existential analogy systems do not seem to lead to general properties for the 'special values'.

At the end of both sections it will be indicated how different widths for the predicates can be included in these systems. In the final section it will be indicated how both types of analogy can be integrated into one system.

2. GENERAL BACKGROUND

Systems

In earlier publications we have already argued for, and made use of, set-theoretic formulations of the 'pure' systems behind the intended application to a monadic predicate language.[2] We will use this convenient way of formulating throughout this paper. Only some terms will hint at the linguistic application.

A regular consistent probability pattern or, simply, a *system* is a tuple

$\langle K, p\rangle$ of which K is a set of so-called (Q-)predicates ($2 \leq |K| < \infty$) and p is a function from K, K^2, K^3 ... to the open interval (0,1) such that

$$\text{for all n} \qquad \sum_{e_n \in K^n} p(e_n) = 1$$

$$\text{for all } n \text{ and } e_n \in K^n \qquad p(e_n) = \sum_{Q_i \in K} p(e_n Q_i)$$

A system gives rise to a number of other probability assignments, such as

$$p(Q_i/e_n) = p(e_n Q_i)/p(e_n) \qquad \text{(special values)}$$

$$p(E_n) = \sum_{e_n \in E_n} p(e_n) \qquad E_n \subseteq K^n$$

A sequence e_n can be seen as the evidence on the basis of the first n trials of experiments. The number of occurrences of Q_i in e_n is indicated by n_i (e_n), the set of instantiated predicates according to e_n (i.e. those Q_i for which $n_i\,(e_n) > 0$) is indicated by $M(e_n)$. If e_n is fixed in the context, which will usually be the case, we simply write n_i and M.

Capital letters W,V (and K,M) indicate non-empty subsets of K, their size is indicated by the corresponding small letters, e.g. $|W| = w$. If k occurs as an index (Q_k) it does not, however, indicate the size of K.

A constituent H_W is the set of infinite sequences in which all of, and only, the members of W occur.[3] The constituents constitute a partition of K^∞, they represent the basic universal generalizations (with existential impact).

Due to the extension theorem of Kolmogorov any system can be decomposed in a *prior distribution*

$$p(H_W) \geq 0 \qquad \sum_{W \subseteq K} p(H_W) = 1$$

and, if $p(H_W) > 0$, *conditional systems* $\langle W, p_W\rangle$ defined by

$$p_W(e_n) = p(H_W \cap e_n\, WW \ldots)/p(H_W) \qquad e_n \in W^n$$

such that

$$p(e_n) = \sum_{W \supseteq M} p(H_W)\, p_W(e_n)$$

$p(H_W/e_n) = p(H_W)\, p_W(e_n)/p(e_n) \qquad W \supseteq M$ (posterior distribution)

$$p(Q_i/e_n) = \sum_{W \supseteq M + Q_i} p(H_W/e_n)\, p_W(Q_i/e_n)$$

(where $M + Q_i$ indicates $M \cup \{Q_i\}$, a device which will be used frequently).

A system is called *open* if $p(H_W) > 0$ for all W, and *closed* if $p(H_K) = 1$ (and hence $p(H_W) = 0$ for $W \neq K$). We will ignore intermediate cases. Note that the conditional systems of an open sysem are closed, i.e. $p_W(H_W) = 1$.

Properties

Apart from analogy properties, the following properties that systems may or may not have are of importance.

> *Order indifference* (strong version): $p(e_n) = p(e'_n)$ if e'_n is an order permutation of e_n.
> *Initial symmetry*: $p(Q_i) = 1/k$ for all i

The following properties are called inductive properties.

> *Instantial confirmation*: $p(Q_i/e_nQ_i) > p(Q_i/e_n)$
>
> *Instantial convergence*: $p(Q_i/e_n) \rightarrow n_i/n \qquad$ if $n \rightarrow \infty$
>
> *Universal confirmation*: $p(H_M/e_nM) > p(H_M/e_n) \qquad M \neq K$
>
> *Universal convergence*: $p(H_M/e_n) \rightarrow 1 \qquad$ if $n \rightarrow \infty$ and $M (\neq K)$ remains constant.

It is evident that a closed system cannot have one of the universal inductive properties.

Widths and distances[4]

A (regular, normalized) *width function* is a real-valued function $\gamma(Q_i) = \gamma_i$ such that $\gamma_i > 0$ and $\Sigma_i\, \gamma_i = 1$.

For both types of analogy the main exposition will be based on the assumption that the widths are equal, i.e. $\gamma_i = 1/k$. In both cases we will finally indicate how to integrate different widths.

For closed systems it is plausible that the equal-width-assumption should lead to initial symmetry. This applies of course also to the closed conditional systems of open systems, i.e. $p_W(Q_i) = 1/w$ $(Q_i \in W)$. Note that this does not imply, nor exclude, that open systems themselves have initial symmetry.

Analogy by similarity is considered to be based on a *distance function* between the predicates, i.e. a real-valued function $d(Q_i, Q_j) \geq 0$ such that $d(Q_i, Q_j) = 0$ if and only if $i = j$, $d(Q_i, Q_j) = d(Q_j, Q_i)$ and $d(Q_i, Q_j) \leq d(Q_i, Q_k) + d(Q_k, Q_j)$.

From time to time we will refer to the corresponding *similarity function* $s(Q_i, Q_j) =_{\text{df}} 1/d(Q_i, Q_j)$. Note that s (Q_i, Q_i) is infinite.

Some defined measures will also be of help: the distance between a predicate and a set of predicates

$$D(Q_i, W) = \sum_{Q_j \in W} d(Q_i, Q_j)$$

and the sum-distance of a set of predicates

$$SD(W) = \sum_{Q_i \in W} \sum_{Q_j \in W} d(Q_i, Q_j)/2$$

It is easy to derive that

$$SD(W + Q_i) = SD(W) + D(Q_i, W) \qquad Q_i \notin W$$

A distance function will be said to be (quantitatively) *isomorphic* if for all Q_i and Q_j there is a 1-1-function f from K-Q_i to K-Q_j such that, for all $Q_k \in K$-Q_i, $d(Q_i, Q_k) = d(Q_i, f(Q_k))$. Intuitively, it means that all predicates have the same 'predicate-environment' as far as distances are concerned.

For an isomorphic distance function one may think of the predicates as the vertices of a square ($k = 4$) or a cube ($k = 8$) or as k arcs of equal length covering the circumference of a circle. A typically non-isomorphic distance function arises if the predicates are represented as a connected sequence of k intervals of real numbers, even if the intervals have equal length.

As is clear from these examples, it is difficult to think of an isomorphic distance function without assuming equal widths. However this may be, it is plausible that in case of equal widths *and* an isomorphic distance function even an open system should have initial symmetry.

It is already intuitively clear that analogy influences can only occur in a degenerate way if $k = 2$. Hence we assume $3 \leq k < \infty$ throughout.

3. INSTANTIAL ANALOGY AS VIRTUAL ANALOGY

The general idea of instantial analogy is that $p(Q_i/e_n)$ may not only depend on n and n_i but also on the 'acquired analogy (by similarity)' on the basis of the particular (number of) occurrences in e_n of the other predicates.

We start by drawing attention to the well-known fact that the special values of a Carnap-system can be written as the weighted mean of an *empirical* factor (n_i/n) and a *logical* factor ($1/k$):

$$pc(Q_i/e_n) = \frac{n_i + \lambda/k}{n + \lambda}$$

$$= \frac{n}{n+\lambda} \cdot \frac{n_i}{n} + \frac{\lambda}{n+\lambda} \cdot \frac{1}{k} \qquad (0 < \lambda < \infty)$$

Note that the weights only depend on n (and sysem constants), that they are non-negative and that their sum equals 1. Moreover, the weight of the logical factor decreases from 1 to 0, whereas the weight of the empirical factor increases from 0 to 1. In the extreme case of $\lambda = 0$ the weight of the empirical factor is 1 for all $n > 0$. That system, with special values n_i/n, is called the *straight rule*.

For completeness we mention the main properties of *C*-systems: instantial confirmation and convergence, initial symmetry, order indifference, and closedness. Only the proof of the last property is somewhat complicated.

Now we can include, in addition, an *analogy factor* in more or less the same way as the logical factor was included:

$$(1) \qquad p(Q_i/e_n) = (n_i + \alpha_i(e_n) + \lambda/k)/(n + \alpha(n) + \lambda) \quad (0 < \lambda < \infty)$$

where the right hand expression can be written as

$$\frac{n}{n + \alpha(n) + \lambda} \cdot \frac{n_i}{n} + \frac{\alpha(n)}{n + \alpha(n) + \lambda} \cdot \frac{\alpha_i(e_n)}{\alpha(n)}$$
$$+ \frac{\lambda}{n + \alpha(n) + \lambda} \cdot \frac{1}{k}$$

To get a proper system we have to assume in addition

(2) $\alpha_i(e_n) \geq 0 \qquad \Sigma_i\alpha_i(e_n) = \alpha(n)$

or, in terms of the analogy factors $\alpha_i(e_n)/\alpha(n)$, that they are non-negative and add up to 1. Note that this also holds for the other factors.

From (1) and (2) it follows now that the weights of the three factors only depend on n (and system constants), that they are non-negative and add up to 1.

A second way of introducing C-systems will help us to proceed.[5] A special value of a C-system can be interpreted as the result of an application of the straight rule to the *real* (empirical) instances of Q_i according to e_n, i.e. n_i and a fixed number of *virtual logical* (VL-) instances of Q_i, i.e. λ/k, for their total sum equals $n + \lambda$. Because this is only a conceptual construction we do not need to assume, here and in what follows, natural numbers of virtual instances, but only non-negative numbers.

To get analogy into a system we not only introduce fixed numbers of virtual logical instances from the start, but also, gradually, *virtual analogy* (VA-) instances, to be distributed over the predicates in a way which will be specified later. From this perspective $\alpha_i(e_n)$ becomes the number of VA-instances acquired by Q_i on the basis of e_n, and $\alpha(n)$ the total number of VA-instances distributed after n trials, to be called the *analogy in* (the first) *n trials*. Of course we assume $\alpha(0) = 0$. The result is that (1) can now be seen as the conceptual application of the straight rule on the sum of the empirical instances, the virtual logical instances and the virtual analogy instances of Q_i on the basis of e_n.

Let us call the number of VA-instances distributed after the n-th trial, i.e.

$$\beta(n) =_{\mathrm{df}} \alpha(n) - \alpha(n-1) \qquad (n > 0)$$

the *marginal analogy* of the n-th trial. In order to get instantial convergence the analogy influence should gradually vanish. Hence it is plausible to assume that $\beta(n)$ decreases from its (finite) initial value $\beta(1) =_{\mathrm{df}} \beta > 0$ to

0. Although $\alpha(n)$ grows, according to this assumption, from 0 to some, possibly infinite, limit value $\alpha(\infty) =_{\mathrm{df}} \alpha$, called the *total analogy of the system*, we have nevertheless the important fact that $\alpha(n)/n \to 0$, even if $\alpha = \infty$.[6]

Some nice properties are the result. As in C-systems, the weight of the logical factor decreases from 1 to 0, whereas the weight of the empirical factor increases from 0 to 1. Moreover, the weight of the analogy factor first increases from 0 (for $n = 0$) to a certain maximum (possibly already for $n = 1$) and then decreases to 0. The system itself has initial symmetry and instantial convergence. All these properties are easy to prove.

The system is also closed.

Proof: Note first that it is sufficient to prove $p(W^n) \to 0$ for all $W \neq K$. *Step* 1. For $e_n \in W^n$, $p(W/e_n)$ is, according to (1) and (2), equal to $(n + \sum_{Q_i \in W} \alpha_i(e_n) + (w/k)\lambda)/(n + \alpha(n) + \lambda) \leq (n + \alpha(n) + (w/k)\lambda)/(n + \alpha(n) + \lambda)$. *Step* 2. $p(W^{n+1}) = \sum_{e_n \in W^n} \sum_{Q_i \in W} p(e_n Q_i) = \sum_{e_n \in W^n} p(e_n) \sum_{Q_i \in W} p(Q_i/e_n) = \sum_{e_n \in W^n} p(e_n)\, p(W/e_n) \leq (n + \alpha(n) + (w/k)\lambda)/(n + \alpha(n) + \lambda) \cdot p(W^n)$. *Step* 3. $p(W^n) \leq \prod_{s=0} (s + \alpha(s) + (w/k)\lambda)/(s + \alpha(s) + \lambda) = \prod_{s=0}^{n-1} (1 - \left(\frac{k-w}{k}\right) \lambda/(s + \alpha(s) + \lambda))$. It is well-known that a product such as the last one goes to 0 if and only if the series $\sum_{s=0}^{n-1} \left(\frac{k-w}{k}\right) \lambda/(s + \alpha(s) + \lambda)$ diverges. The latter is easily seen to be the case if $\alpha(s)/s$ is bounded, for then the terms become comparable to $1/s$. From our assumptions if follows immediately that $\alpha(s) \leq \beta s$ and hence that $\alpha(s)/s$ is bounded by β. *Q.E.D.*

If we add the assumption $\beta < 1$ it also follows that the weight of the empirical factor is larger than that of the analogy factor for $n > 0$, even in such a way that the system gets instantial confirmation.

For easy reference we summarize the assumptions about $\alpha(n)$ and $\beta(n)$

(3)
$$1 > \beta =_{\mathrm{df}} \beta(1) > \beta(n) > \beta(n+1) > 0 \qquad (n > 1)$$
$$\beta(n) \to 0$$
$$\alpha(0) = 0,\ \alpha(n) =_{\mathrm{df}} \beta(1) + \beta(2) \ldots \beta(n) \qquad (n \geq 1)$$
$$\alpha(n) \to \alpha(\infty) =_{\mathrm{df}} \alpha \leq \infty$$

There are of course many different ways to specify $\beta(n)$ in accordance with (3). A simple example is $\beta(n) = 1/(n+1)$, leading to infinite total analogy. Note that $\beta(n) = 1/n$ would not guarantee instantial confirmation for $n = 0$, i.e. $p(Q_i/Q_i) > p(Q_i)$ is not yet guaranteed.

A particularly attractive way of specifying $\beta(n)$ is to assume that $\beta(n+1)$ is a constant fraction of $\beta(n)$, i.e. to assume that there is a real number x such that

(X) $\quad \beta(n+1) = x\beta(n) \qquad (0 < x < 1)$

This proposal leads directly to $\beta(n+1) = x^n\beta$ and to $\alpha(n) = \beta(1\text{-}x^n)/(1\text{-}x)$, and hence $\alpha = \beta/(1\text{-}x)$. Here, x might be called the analogy rate and $(1\text{-}x)$ the analogy decay-rate (note that $\beta(n)\text{-}\beta(n+1) = (1\text{-}x)\,\beta(n)$). Of course, starting from a certain (finite) value for the total analogy of the system, we may either prefer high initial analogy (β) and high decay-rate or low initial analogy and low decay-rate. Although (X) is rather attractive, what follows will in no way depend on (X), but only on (3).

Our next task is to regulate the distribution of the marginal analogy over the predicates. In particular, we have to specify the *marginal analogy profit* of Q_i from the occurrence of Q_j after e_n, i.e. $\alpha_i(e_nQ_j)\text{-}\alpha_i(e_n)$, for all i,j and e_n and in such a way that their sum over i equals $\beta(n+1)$.

For this purpose we will assume *constant* marginal analogy profit *rates*, i.e. we will assume that there is an *analogy matrix* of constants $a_i(j)$ such that

(4) $\quad \alpha_i(e_nQ_j)\text{-}\alpha_i(e_n) = a_i(j)\,\beta(n+1)$

It results from the plausible step-wise assumptions that the marginal analogy profit of Q_i from Q_j does not depend on the already acquired profit of Q_i from e_n, and hence can be written as $\beta_i(j,n)$, and that the corresponding rate $\beta_i(j,n)/\beta(n+1)$ does not depend on n.

From (2), (3) and (4) it follows that the analogy matrix is such that

(5.1) $\quad 0 \leq a_i(j)$
(5.2) $\quad \Sigma_i a_i(j) = 1$

We add to these conditions

(5.3) $\quad a_i(i) = 0$

in order to exclude that a predicate has analogy profit from itself. Note that (1) already guarantees that a predicate has (constant) 'instantial profit' from itself.

It is easy to check that substitution of (4) in (1), using (5.3), leads to

$$(6.1) \qquad p(Q_i/e_nQ_i) = \frac{n_i + 1 + \alpha_i(e_n) + \lambda/k}{n + 1 + \alpha(n) + \beta(n+1) + \lambda}$$

$$(6.2) \qquad p(Q_i/e_nQ_j) = \frac{n_i + \alpha_i(e_n) + a_i(j)\,\beta(n+1) + \lambda/k}{n + 1 + \alpha(n) + \beta(n+1) + \lambda}\; i \neq j$$

Conversely, if we add initial symmetry

$$(6.3) \qquad p(Q_i) = 1/k$$

the whole system is determined by (3), (5) and (6). Such systems will be called *potential virtual analogy* systems (VAp-systems). Besides initial symmetry, VAp-systems have the properties of instantial confirmation and convergence; moreover, they are closed.

Our final task is to relate the analogy matrix to the presupposed distance function in a plausible way. But let us first note that VAp-systems have already an important analogy property. Due to (3) and (5) the instantial profit which a predicate gets from itself is always larger than the analogy profit it could get from any other predicate (for $1 > a_i(j)\,\beta(n+1)$). Consequently, VAp-systems have

self-similarity: $p(Q_i/e_nQ_i) > p(Q_i/e_nQ_j) \qquad i \neq j$

of which the plausibility directly follows from the fact that

$$0 = d(Q_i, Q_i) < d(Q_i, Q_j) \qquad \mathrm{i} \neq \mathrm{j}$$

The main problem of finding a plausible general principle of analogy probably stems from the fact that at first sight the following principle seems to be the obvious generalization of self-similarity:

Carnap-analogy: if $d(Q_i, Q_j) </= d(Q_i, Q_k)$ then $p(Q_i/e_nQ_j) >/= p(Q_i/e_nQ_k)$

From (6) we may immediately infer for all VAp-systems:

(7) if $a_i(j) >/= a_i(k)$ then $p(Q_i/e_nQ_j) >/= p(Q_i/e_nQ_k)$

Hence, VAp-systems have Carnap-analogy if we add the assignment rule

$$\text{(C)} \qquad \text{if } d(Q_i, Q_j) </= d(Q_i, Q_k) \text{ then } a_i(j) >/= a_i(k) \qquad i \neq j$$

of which the inequality-part reads, in appealing terms, 'if Q_j is closer to Q_i than Q_k, then Q_i profits more from Q_j than from Q_k'.

Carnap was well aware that his principle was not entirely plausible and he introduced a number of weaker versions[9]. From our approach it is immediately clear that (C), and hence Carnap-analogy, is in general implausible. The reason is that the 'predicate-environment' of Q_j may differ strongly from that of Q_k. Hence, the way in which Q_j and Q_k have to distribute their analogy influence in accordance with their respective environments may differ so much that $a_i(j) < a_i(k)$, contrary to what (C) prescribes.

Of course, if we exclude different environments, this possibility should also be excluded. In Section 2 we have already specified the formal interpretation of the idea that all predicates have the same environment: the distance function is *isomorphic*. An isomorphic distance function, however, also suggests directly that the analogy matrix should be symmetric in this case, i.e. $a_i(j) = a_j(i)$. In view of this we will return to this case later on.

In contrast to (C) there is, given the way in which we have constructed VAp-systems, an entirely plausible assignment rule:

$$\text{(V)} \qquad \text{if } d(Q_i, Q_j) </= d(Q_i, Q_k) \text{ then } a_j(i) >/= a_k(i) \qquad i \neq j$$

of which the inequality-part reads, again in appealing but now also convincing terms, 'if Q_j is closer to Q_i than Q_k then Q_j profits more from Q_i than Q_k'.

Although the verbal formulation is rather plausible, it is not directly clear to what general principle of analogy it gives rise. However, from its formal version (V) and (6) (interchanging Q_i and Q_j) we see immediately that the principle looked for, and guaranteed by (V), is:

$$\textit{virtual analogy}\text{: if } d(Q_i, Q_j) </= d(Q_i, Q_k) \text{ then}$$
$$p(Q_j/e_nQ_j)\text{-}p(Q_j/e_nQ_i) </= p(Q_k/e_nQ_k)\text{-}p\ (Q_k/e_nQ_i)$$

Note first that virtual analogy, like Carnap-analogy, implies self-similarity (substitute in both cases $i = j$), which was already satisfied without (V) (or (C)). According to self-similarity the difference $p(Q_j/e_nQ_j)$-p (Q_j/e_nQ_i) is always positive (for $i \neq j$), hence we might call this difference the Q_i-

substitution-loss for Q_j (after e_n). In these terms, virtual analogy adds to self-similarity that the Q_i-substitution-loss is the larger for a predicate the larger its distance from Q_i.

VA*p*-systems satisfying (V) will be called VA-systems. They have the property of virtual analogy.

From (V) it follows directly that, if the analogy matrix is symmetric ($a_i(j) = a_j(i)$), (C) is also satisfied and hence (V) guarantees in this case Carnap-analogy. As we have already remarked, a symmetric matrix is plausible if the distance function is isomorphic. Hence, if we accept this as a general rule, a VA-system based on an isomorphic distance function has virtual analogy as well as Carnap-analogy.

The foregoing can be illustrated by a plausible quantitative link between the analogy matrix and the distance function or, more conveniently, the similarity function $s(Q_i, Q_j)$ ($=_{df} 1/d(Q_i, Q_j)$):

$$\text{(Vq)} \qquad a_i(i) = 0,\ a_j(i) = s(Q_i, Q_j)/\sum_{k \neq i} s(Q_i, Q_k) \qquad i \neq j.$$

It is easy to check that (Vq) is in agreement with the general requirements (5) for the analogy matrix and that it implies (V). Moreover, it is easily seen to lead only to a symmetric matrix if $\sum_{k \neq i} s(Q_i, Q_k) = \sum_{k \neq j} s(Q_j, Q_k)$ for all i and j. Apart from possible 'accidental' exceptions this condition is only satisfied if the distance function is isomorphic.

VA-systems have still other analogy properties than virtual analogy. In view of the next section about existential analogy the following property, which is easy to prove, is particularly interesting:

(VEA) *virtual existential analogy*: if e_n ($n > 0$) is such that $n_i = n_j = 0$ and if, for all Q_k for which $n_k > 0$ $d(Q_i, Q_k) \leq d(Q_j, Q_k)$, with at least one proper inequality, then $p(Q_i/e_n) > p(Q_j/e_n)$

It is not difficult to check that VAp-systems satisfying (C) have this property only if the analogy matrix is symmetric, hence, only if (V) is also satisfied.

Let us now direct our attention to the possibility of generalizing VA-systems in such a way that they include different widths. It is well-known that C-systems can be generalized (to so-called GC-systems) by replacing the logical factor $1/k$ by γ_i, or similarly by replacing the virtual logical instances λ/k by $\gamma_i\lambda$. The systems so obtained have special values $p(Q_i/e_n) = (n_i + \gamma_i\lambda)/(n + \lambda)$. Moreover, they have all the properties of C-systems (instantial confirmation and convergence, order indifference, closedness) except of course initial symmetry, which is only satisfied for the sub-class of C-systems. It is easy to check that GC-systems have self-similarity, but apart from this they do not have virtual analogy because the substitution--loss is the same for all combinations of predicates (viz. $1/(n + 1 + \lambda)$). They also lack, trivially, Carnap-analogy.

It is no problem to see that we can generalize the introduction of VA-systems in the same way by replacing the term λ/k in the numerators of (1) and (6) by $\gamma_i\lambda$. The resulting GVA-systems have all the properties of VA-systems, again except of course initial symmetry, which holds only for the subclass of VA-systems. Moreover, GVA-systems have Carnap-analogy if the distance function is isomorphic. But it is doubtful whether there are (interesting) cases of this, when the widths are not equal.

From the foregoing we conclude that GVA-systems include analogy influences in Carnapian spirit, with two important deviations. The first is that Carnap-analogy is not the general analogy property but virtual analogy. The second deviation is that GVA-systems do not have order indifference. This is indeed an important deviation, for Carnap has always considered order indifference as one of the basic axioms of (pure) inductive probability. So we will comment on it in some detail.

That GVA-systems lack order indifference is easy to prove by first showing that they even fail to satisfy *both* of its analytic components[8]:

(a) $p(Q_iQ_j/e_n) = p(Q_jQ_i/e_n)$

(b) $p(Q_i/e_n) = p(Q_i/e'_n)$ if e'_n is an order permutation of e_n.

It is important to note that this failure of order indifference has not arisen at some later stage in the construction procedure, but that this failure was already implied by our very first step. That is, even if we include different widths in (1), it follows immediately that the only systems satisfying this liberalized version of (1) and (a) are GC-systems (for $\alpha_i(e_n)$ is forced to be independent of e_n), i.e. systems of which we have seen already that they lack virtual analogy (and Carnap-analogy) as far as this exceeds self-similarity.[9]

Probably the most important reason for wanting to have order indifference is that *order differences* suggest that the order of occurrences of predicates is informative. This is what Carnap apparently had in mind when he discussed briefly the idea of 'analogy by proximity' as opposed to 'analogy by similarity' (see note 1). In the case of analogy by proximity order difference would have a natural background.

However, in our treatment of analogy by similarity we get order differences not because we take them as informative but because they happen to be the consequence of a plausible way of dealing with such analogy influences.

An obvious question nevertheless is whether GVA-systems have some order *convergence* properties. To begin with, it is trivial that $p(e_n')$ goes to $p(e_n)$, if e_n' is an order permutation e_n, for in every system $p(e_n)$ goes to 0 (only with the exception of uniform sequences $Q_iQ_iQ_i\ldots$ in open systems). Moreover, systems which have instantial convergence, and hence GVA-systems, satisfy the respective limit-*re*formulations of (a) and (b). Integrating instantial convergence they read as follows:

(a') $$p(Q_iQ_j/e_n) \rightarrow \frac{n_in_j}{n(n+1)} \leftarrow p(Q_jQ_i/e_n)$$

(b') $$p(Q_i/e_n) \rightarrow \frac{n_i}{n} \leftarrow p(Q_i/e_n') \text{ if } e_n' \text{ is an order permutation of } e_n.$$

These order convergence properties are of crucial relevance for the evaluation of the lack of order indifference. For, if the order of occurrence of the predicates would be considered as really informative about some underlying process, then order differences should not of course disappear in the long run, or at least not as a rule. Hence, in view of the fact that GVA-systems have the order convergence properties (a') and (b') we may conclude that the order differences arising in GVA-systems cannot be interpreted as if we have taken the order of predicates as informative.

In our opinion the lack of order indifference is, in the light of the foregoing, fully justified and there is only one thing to regret. It is well-known that a number of important inductive systems can be introduced in a sophisticated way by assuming in addition to order indifference the relevant arguments for the special values. For example, C-systems can be obtained in this way by assuming that $p(Q_i/e_n)$ only depends on n_i and n. In these approaches order indifference functions as a very strong 'calculation-principle'. It is clear that such a sophisticated way of introduction is impossible for GVA-systems.

As far as we know GVA-systems have not been presented in the literature, let alone the particular way of introduction as described in this paper. However, D. Costantini (1983) has formulated a system which comes close to ours. He assumes a 'natural number distance function' between the predicates and that a predicate has only analogy influence on the two 'neighbouring' predicates. Although his method of introduction is not very clear to me, he arrives at the special values, in our notation:

$$\text{(DC)} \qquad p(Q_i/e_n) = \frac{n_i + a(n_{i-1} + n_{i+1}) + \gamma_i\lambda}{n + 2an + \lambda}$$

with special devices for $i = 1$ and $i = k$, which do not matter here.

It is easy to check that (DC) represents a GVA-like system with $\beta(n) = 2a$ and hence $\alpha(n) = 2an$ and with analogy matrix:

$$a_{i-1}(i) = a_{i+1}(i) = \tfrac{1}{2},\ a_j(i) = 0 \text{ if } j \neq i-1, i+1$$

The analogy matrix is in accordance with our condition (5) and can easily be generalized to include more distant analogy influences. Nevertheless it is not a GVA-system, because $\beta(n)$ is constant, i.e. the marginal analogy is constant. As Costantini is well aware his system lacks, precisely due to this feature, instantial convergence. To rectify this he suggests, apart from an ad hoc proposal, modifying the system by replacing a by a/n in (DC). Although this leads to instantial convergence, it is also not a GVA-system[10] for now it fails to satisfy (4), in particular, the marginal analogy profit of Q_{i-1} and Q_{i+1} from Q_i is not independent of the profit these predicates have already obtained from e_n. Costantini is also not very happy with this suggestion, but for different reasons, which I do not fully understand.

We conclude this section with a sketch of an objective model in which the objective probabilities are precisely those of a GVA-system.

Elsewhere[11] we have shown in some detail that generalized Carnap-systems have such an objective model, viz. a so-called Polya urn-model. Start with an urn containing $\gamma_i\lambda$ 'Q_i-balls', hence in total λ balls. A trial is a random selection of a ball with replacement of the drawn ball *and* the addition of a new ball with the same Q-predicate as the one drawn. It is easy to verify that this procedure leads to a GC-system.

Given the way in which we have introduced VA-systems it is easy to imagine how we have to modify and generalize this model to a GVA-model. Consider a reservoir containing some kilos of (grains of) sand of k different colours. The probability of drawing a Q_i-grain is assumed to be proportional to the total weight of the Q_i-grains in the reservoir. At the start there are for each Q_i $\gamma_i\lambda$ kilo of Q_i-sand. If the n-th trial results in a Q_i-grain (that grain is thrown back and) we add one kilo of Q_i-sand as well as, for all $j \neq i$, $a_j(i)\,\beta(n)$ kilo of Q_j-sand, where $\beta(n)$ is in accordance with (3) and $a_j(i)$ in accordance with (5). This model of course behaves like a GVA-system; the virtual analogy 'instances' are as it were materialized, together with the real instances and the virtual logical instances. Among others, the model makes it intuitively plausible that GVA-systems are closed: the probability is 0 that in the long run one or more types of sand disappear from the reservoir.

4. EXISTENTIAL ANALOGY

The basic idea of existential analogy is that the distances between the predicates are accounted for in the prior distribution of open systems.

The consequences of this procedure will be studied in open systems with related conditional C-systems, viz. there is a real number ρ, $0 < \rho < \infty$, such that the conditional systems $\langle W, p_W\rangle$ are C-systems with 'λ_W' is $w\rho$. Hence, the conditional special values are given by

$$(8) \qquad p_W(Q_i/e_n) = (n_i + \rho)/(n + w\rho) \qquad e_nQ_i \in W^{n+1}$$

Note that the conditional systems have initial symmetry.

An open system of the kind described will be called a *potential* existential analogy system (EAp-system). An EAp-system is a generalization of the so-called K-system, which was introduced by Hintikka and Niiniluoto (1976). To get a K-system we have to add the assumption that $p(H_W)$ only depends on w[12]. It is clear that this additional assumption blocks the possibility of analogy via the prior distribution.

We will first discuss some properties of EAp-systems. From the fact that C-systems have order indifference it follows immediately that EAp-systems also have this property. We have investigated a class of systems elsewhere[13] which trivially includes EAp-systems. From that investigation we may conclude here that EAp-systems have all four inductive properties specified in Section 2, i.e. instantial/universal confirmation/convergence.

From the particular assumption $\lambda_W = w\rho$ it follows

(9) $p_W(Q_i/e_nM)$ $(=_{\text{df}}\ p_W(e_nQ_i)/p_W\ (e_nM)\ ) = (n_i + \rho)/(n + m\rho)$
$Q_i \in M$

and from this we get

(10) $p(Q_i/e_nM) = (n_i + \rho)/(n + m\rho) \qquad Q_i \in M$

and that there is a function $g(M, n)$ such that

(11) $p(M/e_n) = g(M, n)$

In fact EAp-systems could have been introduced by assuming an open system with order indifference such that $p(Q_i/e_n)$ may only depend on Q_i, M and n if $Q_i \notin M$, whereas $p(Q_i/e_nM)$, $Q_i \in M$, may only depend on m, n and n_i[14]. If we would replace these dependencies by '$p(Q_i/e_n)$ may only depend on m, n and n_i' we would get the K-systems.

For our purposes one additional feature of EAp-systems will be important: from (8) it follows directly

(12) $p_W(e_n) = p_V(e_n)$ if $w = v$ and $e_n \in W^n$, $e_n \in V^n$

Already at this stage we can evaluate EAp-systems with respect to the analogy properties discussed in the preceding section. To begin with, EAp-systems have, as far as already instantiated predicates are concerned, self-similarity:

$$\text{if } Q_i,\ Q_j \in M \text{ then } p(Q_i/e_nQ_i) > p(Q_i/e_nQ_j)$$

This follows directly from (10) and (11) for they imply

(13) $p(Q_i/e_nQ_i) = g(M, n + 1)\ (n_i + 1 + \rho)/(n + 1 + m\rho)$
$Q_i \in M$

(14) $p(Q_i/e_nQ_j) = g(m, n + 1)\ (n_i + \rho)/(n + 1 + m\rho)$
$Q_i,\ Q_j \in M$

But they do not have the additional content of virtual analogy nor that of Carnap-analogy as far as already instantiated predicates is concerned (i.e. Q_i, Q_j, $Q_k \in M$). Both claims follow immediately from (13) en (14).

The situation is more complicated if one or more of the relevant predicates is not instantiated in e_n. As a rule the question of whether or not a certain combination of predicates, not all of them instantiated in e_n, has one of the properties depends on the prior distribution[15].

This holds in particular for the property we have called virtual existential analogy (VEA) which is exclusively defined for two non-instantiated predicates (in the present notation):

(VEA) if $Q_i, Q_j \notin M \neq 0$, and for all $Q_k \in M$ $d(Q_i, Q_k) \leq d(Q_j, Q_k)$, with at least one proper inequality, then $p(Q_i/e_n) > p(Q_j/e_n)$

Now we will specify four plausible conditions on the prior distribution of increasing strength. All of them will lead to plausible analogy properties for the posterior distribution, but even the strongest one will not yet enable us to prove (VEA).

(I) If $Q_i, Q_j \notin W$ and for all $Q_k \in W$ $d(Q_i, Q_k) \leq d(Q_j, Q_k)$ (with at least one proper inequality) then $p(H_{W+Q_i}) \geq (>) p(H_{W+Q_j})$

Conditions of this kind are to be read as starting with 'For all W (and V)'. Using (12) it is easy to prove that an EA-I-system (i.e. an EAp-system satisfying (I)) has the property

(EA-I) If $W \supseteq M$ and $Q_i, Q_j \notin W$ and for all $Q_k \in W$ $d(Q_i, Q_k) \leq d(Q_j, Q_k)$ (with at least one proper inequality) then $p(H_{W+Q_i}/e_n) \geq (>) p(H_{W+Q_j}/e_n)$

In our opinion (I) and (EA-I) are highly plausible. Note that the translation of the antecedent of (I) in terms of the similarity function does not create any difficulties.

From a proof attempt it will become clear that it is not possible to prove (VEA) in an EA-I-system. Assume the conditions of (VEA). To be proved is $p(Q_i/e_n) > p(Q_j/e_n)$ and hence

$$\sum_{W \supseteq M+Q_i} p(H_W/e_n)\, p_W(Q_i/e_n) > \sum_{V \supseteq M+Q_j} p(H_V/e_n)\, p_V(Q_i/e_n)$$

By cancelling the common terms and substituting (8) it remains to be proved that

$$\sum_{\substack{W \supseteq M+Q_i \\ Q_j \notin W}} p(H_W/e_n)\, (\rho/(n + w\rho)\,) > \sum_{\substack{V \supseteq M+Q_j \\ Q_i \notin W}} p(H_V/e_n)\, (\rho/(n + v\rho)\,)$$

In view of (I) and (EA-I) it is clear that it is only possible to prove this if we would have for all $W \supseteq M + Q_i$, $Q_j \notin W$, $p(H_W) \geq p(H_{W-Q_i+Q_j})$ with at least one inequality. But the conditions of (VEA) only allow us to infer from (I): $p(H_{M+Q_i}) > p(H_{M+Q_j})$. To put it differently, to get something like (VEA) from (I) we would have to strengthen the conditions of (VEA) in such a way that they are almost never satisfied. Moreover, they would be specific conditions for the combination of Q_i, Q_j and M. Hence, it is unlikely that we could get (VEA) itself by general principles like (I).

Our first strengthening of (I) will make use of the defined notion of the distance between a predicate and a set of predicates.

(II) If $Q_i, Q_j \notin W$ and $D(Q_i, W) </= D(Q_j, W)$ then $p(H_{W+Q_i}) >/= p(H_{W+Q_j})$

Again with (12) it is easy to prove that EA-II-systems satisfy

(EA-II) If $W \supseteq M$ and $Q_i, Q_j \notin W$ and $D(Q_i, W) </= D(Q_j, W)$ then $p(H_{W+Q_i}/e_n) >/= p(H_{W+Q_j}/e_n)$

It is also easy to see that (II) and (EA-II) imply (I) and (EA-I), respectively.

Although (II) and (EA-II) look almost as plausible as their corresponding I-versions, there is one problem with them. Consider the notion of the similarity between a predicate and a set of predicates $S(Q_i, W)$, $Q_i \notin W$, defined as $\sum_{Q_j \in W} s(Q_i, Q_j)$. It is easy to check that for example the condition $D(Q_i, W) < D(Q_j, W)$ does not coincide with the condition $S(Q_i, W) > S(Q_j, W)$. Nevertheless, (II) and (EA-II) would seem to be equally plausible if we had used in it the condition $S(Q_i, W) >/= S(Q_j, W)$. Hence, we are confronted with two equally plausible conditions, (II) and the indicated variant, which leave room for straightforward conflicting prescriptions. For the simple reason that we are more used to thinking in terms of distances than in terms of similarities we will choose (II) instead of the discussed alternative. A similar choice is involved in the stronger versions which follow.

The next strengthening of (II) is, apart from the preceding qualification, again rather plausible. The sum-distance SD(W) is, at least for sets of the same size, a measure of the heterogeneity of a set of predicates. This observation suggests:

(III) If $SD(W) </= SD(W')$ and $w = w'$ then $p(H_W) >/= p(H_{W'})$

with the direct consequence from (12)

(EA-III) If $W, W' \supseteq M$ and $SD(W) </= SD(W')$ and $w = w'$ then $p(H_W/e_n) >/= p(H_{W'}/e_n)$

Again it is easy to check that (III) and (EA-III) imply (II) and (EA-II), respectively (use $SD(W + Q_i) = SD(W) + D(Q_i, W)$ if $Q_i \notin W$).

Our last strengthening will be the specification of a quantitative relation between the distance function and the prior distribution. There are of course $w(w\text{-}1)/2$ different combinations of two different predicates in set W. Hence, $2SD(W)/(w(w\text{-}1))$ indicates the average distance between two different predicates in W, which is undefined for one-element sets. This average distance might well be called the *heterogeneity* of W (Het(W)) and 1/Het(W) the *homogeneity* of W (Hom(W)), which is also undefined for one-element sets. Let us agree that one-element sets have all the same prior probability (q) ($0 < kq < 1$) for the corresponding constituent, i.e. $p(H_{Q_i}) = q$ for all Q_i. Now it is an obvious idea to specify the prior distribution in such a way that $p(H_W)$, as far as sets with $w \geq$ are concerned, becomes proportional to the homogeneity of W. This idea leads to:

$$\text{(IV)} \qquad p(H_W) = \frac{\text{Hom}(W)}{\sum\limits_{2 \leq v \leq k} \text{Hom}(V)}(1\text{-}kq) \qquad (w \geq 2)$$

$$p(H_{Q_i}) = q \qquad 0 < kq < 1$$

To derive the most general consequence for the posterior distribution we specify a property of EAp-systems in addition to (12):

(12′) $p_W(e_n) > p_V(e_n)$ if $w < v$ and $e_n \in W^n$ and $e_n \in V^n$

With the help of (12) and (12′) it is not difficult to prove that EA-IV-systems satisfy

(EA-IV) If $W, V \supseteq M$ and $2 \leq w \leq v$ and $\text{Hom}(W) \geq \text{Hom}(V)$ then $p(H_W/e_n) \geq p(H_V/e_n)$. If, in addition, $w < v$ *and/or* $\text{Hom}(W) > \text{Hom}(V)$ then $p(H_W/e_n) > p(H_V/e_n)$.

Again (IV) and (EA-IV) are rather plausible and they imply (III) and

(EA-III), respectively. We would get variants if we had defined the homogeneity directly as the average similarity between two different predicates.

It is not difficult to check that the additional content of (IV) compared with (I) does not add anything of the kind which would be needed to complete the proof attempt of (VEA) for EA-IV-systems. Although (I)-(IV) do not exhaust all possibilities, the conclusion is tempting that (VEA) is a property which cannot be obtained by special restrictions on the prior distribution.

A question left untouched thus far is under what conditions EA-systems have initial symmetry. Although the conditional systems have the property, it is clear that this does not guarantee that the system itself has it. Using (8), initial symmetry is seen to hold if and only if

$$(15) \qquad p(Q_i) = \sum_{Q_i \in W} p(H_W)\,(1/w) = 1/k$$

As in the case of the preceding section this property is especially to be expected when the distance function is isomorphic. It is not difficult to see that (III) is just strong enough to guarantee initial symmetry for an isomorphic distance function. For, such a function implies that, for fixed Q_i and Q_j, there corresponds to every W, $Q_i \in W$, a unique set W' with $w' = w$, $Q_j \in W'$ and $\mathrm{SD}(W) = \mathrm{SD}(W')$ and hence, by (III), $p(H_W) = p(H_{W'})$, which leads directly to (15).

Inspection of (IV) shows that even this condition does not exclude the possibility of 'accidental' initial symmetry in case of a non-isomorphic distance function.

There is an obvious way to include widths in our treatment, viz. by assuming open systems with related GC-systems as conditional systems. Let γ_i be the 'basic' width function and define γ_W as $\sum_{Q_i \in W} \gamma_i$. Replace (8) now by

$$(8G) \qquad p_W(Q_i/e_n) = (n_i + \gamma_i\lambda)/(n + \gamma_W\lambda) \qquad Q_i \in M$$

Note that the width of Q_i 'within W' now corresponds to γ_i/γ_W and 'λ_W' to $\gamma_W\lambda$. Note also that we get back (8) if we substitute $\gamma_i = 1/k$ and $\lambda = k\rho$. For the thus construed GEAp-systems the expressions corresponding to (9) and (10) are

(9/10G) $p_W(Q_i/e_nM) = p(Q_i/e_nM) = (n_i + \gamma_i\lambda)/(n + \gamma_M\lambda) \quad Q_i \in M$

whereas (11) remains the same.

Again it holds that such systems have order indifference and that they satisfy instantial/universal confirmation/convergence. And also, that they lack the properties of virtual analogy and Carnap analogy (at least) for already instantiated predicates.

Moreover, it is clear that the conditions (I)-(IV) do not lose their plausibility. However, it is no longer the case that (12) holds in general. Hence, the properties for the posterior distribution (EA-I–EA-IV) can no longer be proved. But this is not surprising and it may well be that restricted versions can be proved.

Finally, nothing will change with respect to the unprovability of (VEA). From the fact that EA-systems and GEA-systems do not have this property, it should not, of course, be concluded that existantial analogy, as exemplified by (I)-(IV), has at the most consequences for the posterior distribution. On the contrary, (I)-(IV) have strong effects on the special values $p(Q_i/e_n)$, but we have not succeeded in finding nice properties describing such effects.

5. CONCLUDING REMARKS

In Sections 3 and 4 we have seen that there are two quite different ways to design inductive systems with analogy influences based on a distance function. In case of virtual analogy, predicates acquire analogy influences on the basis of experimental results. In case of existential analogy all analogy influences are a priori stored in the prior distribution.

It must be admitted that the relation between the two notions of analogy has not been made clear in this paper. The main task in this respect is to find systematic analogy effects on the special values due to existential analogy. Further research on this point will be especially desirable for those who welcome systems with existential analogy because of the retention of order indifference, as opposed to systems with virtual analogy.

Other research may be directed along the following lines. First, whereas closed systems cannot have existential analogy, open systems with virtual analogy can of course be designed: open systems with VA-systems as conditional systems and $p(H_W)$ depending only on w.

Second, it is also obvious that it is possible to construct systems in which both types of analogy are integrated: open systems with VA-systems as conditional systems *and* a prior distribution satisfying one of the conditions (I)-(IV).

We will not study here the precise properties of the two indicated kinds of open systems, but restrict ourselves to three remarks.

1. Both kinds of open systems fail to have order indifference.

2. The systems developped by Niiniluoto (1981) and Spohn (1981), being open systems, evidently without existential analogy, should be compared with the first-mentioned type of open systems with virtual analogy. In the introduction we mentioned already that Spohn (1981) pointed out that the two proposals of Niiniluoto are not acceptable. Hence, it remains to carry out the comparison with Spohn's own proposal. However, in the light of the ad hoc character of that proposal, this comparison seems to be neither an easy nor an important task.

3. As to the value of this kind of research I must concede that my scepticism about the relevance for the philosophy of science, as described at the end of my (1978) and (1981), has not decreased. Fortunately, however, the relevance for statistics seems to increase, for it is clear from the present paper that, in particular, virtual analogy ideas might well find application in statistics.

NOTES

* The author would like to thank the Netherlands Institute for Advanced Study (Wassenaar) for making the present research possible, and Anne Simpson (NIAS) for improving the English.

[1] In Section 17.C Carnap pays also some attention to 'analogy by proximity', where the order of outcomes is interpreted as informative.

[2] Kuipers (1978), (1980), (1981). Some of the set-theoretic notations will be informal.

[3] Constituents can be introduced formally by the definitions $H_W(n) = \{e_n \in K^n/M(e_n) = W\}$ and $H_W = \bigcup_{n=w}^{\infty} H_W(n)WWW\ldots$.

[4] For general considerations about widths and distances see Carnap (1980), Section 14.

[5] According to a letter from R. Jeffrey, Carnap has also thought of this way of introducting C-systems. As far as Jeffrey and I know Carnap has never mentioned it in his publications.

[6] This is a direct consequence of Cauchy's general limit theorem for arithmetic means of convergent sequences, telling that the arithmetic means converge to the same limit as the convergent sequence itself.

[7] See Carnap (1980), pp. 46-47. Our definition of Carnap-analogy corresponds to his version B_2.

[8] The non-trivial fact that (a) and (b) together imply order indifference is proved in Kuipers (1978), pp. 40-41.

[9] As far as the violation of 'indifference with respect to the past', i.e. (b), is concerned, the crucial step was made by (4). Although (4) is in our opinion highly plausible, it may be of interest to study the system that arises if we replace (4) by

$$(4') \qquad \alpha_i(e_n) = \frac{\alpha(n)}{n} \Sigma_j\, n_j a_i(j)$$

which is easily seen to satisfy (b). As Costantini rightly remarked in a personal letter, (b) may be desirable for practical reasons: "Almost all statistical data do not take into account the order in which predicates are observed".
[10] In fact there arises a special case of an extreme case of (4'), see note 9, i.e. α(n) is constant, say α:

$$(4'') \qquad \alpha_i(e_n) = \frac{\alpha}{\mathrm{n}} \Sigma_j n_j a_i(j)$$

Note that constant $\alpha(n)$ is formally excluded by (3), for (3) requires that $\alpha(n)$ grows. But in the set-up of (4') this is no longer a necessary requirement. In fact one may argue that (4") is the natural special case of (4'), with a plausible interpretation: there is a fixed number α of VA-instances, which are redistributed after each trial in accordance with the observed relative frequencies and the analogy matrix. A full study of (4") may be worthwhile. However, one objection is easy to make. It is not difficult to check that self-similarity is only guaranteed if $\alpha < 1$. If we combine this with the fact that the weight of the analogy factor becomes $\alpha/(n + \alpha + \lambda)$, it turns out that this weight is always smaller than $1/(n + 1 + \lambda)$, i.e. this weight is doomed to decrease more rapidly than seems reasonable.
[11] See Kuipers (1978), Section 5.3.3.
[12] That the class of systems obtained in this way is not only a subclass of K-systems but actually coincides with the class of K-systems is proved in Kuipers (1978), Section 6.6.
[13] See Kuipers (1978), Section 6.4. and 6.11. The more general class referred to is called the class of GH-systems.
[14] This is in fact a generalization of the way in which Hintikka and Niiniluoto (1976) introduced K-systems. For more details about the described introduction of EAp-systems see Kuipers (1981), where they are named GSH$_a$-systems.
[15] In Kuipers (1981) we have claimed that GSH-systems, which trivially include EAp-systems, have the property of self-similarity in general. Unfortunately, the proof-sketch of this claim in note 7 of that paper does not work for the case $Q_i \in M$, $Q_j \notin M$.

REFERENCES

Carnap, R.: 1980, 'A Basic System of Inductive Logic, Part 2', in Jeffrey (ed.) (1980), pp. 7-155.

Costantini, D.: 1983, 'Analogy by Similarity', *Erkenntnis* **20**, 103–114.

Hintikka, J. and Niiniluoto, N.: 1976, 'An Axiomatic Foundation for the Logic of Inductive Generalization', *Formal Methods in the Methodology of Empirical Sciences*, eds. M. Przełecki, K. Szaniawski, R. Wójcicki, D. Reidel, Dordrecht, pp. 57-81. Reprinted in Jeffrey (ed.) (1980), pp. 157–181.

Jeffrey, R. (ed.): 1980, *Studies in Inductive Logic and Probability*, Vol. II, University of California Press.

Kuipers, T.: 1978, *Studies in Inductive Probability and Rational Expectation*, D. Reidel, Dordrecht.

Kuipers, T.: 1980, 'A Survey of Inductive Systems', in Jeffrey (ed.) (1980), pp. 183–192.

Kuipers, T.: 1981, 'What Remains of Carnap's Program Today?', to appear in *Probability, Statistics and Inductive Logic*, Proc. Int. Conf. Luino, 1981, special issue of *Epistemologia*, eds. D. Costantini, E. Agazzi.

Niiniluoto, I.: 1981, 'Analogy and Inductive Logic', *Erkenntnis* **16**, 1–34.

Spohn, W.: 1981, 'Analogy and Inductive Logic: A Note on Niiniluoto', *Erkenntnis* **16**, 35–52.

Manuscript submitted 27 May 1983
Final version received 26 October 1983

Dept. of Philosophy
University of Groningen
Westersingel 19
9718 CA Groningen
Holland